ECL

AD-A229 262

DTIC FILE COPY

DTIC ELECTE NOV 16 1990 S B D

(2)

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188
1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) UCD Statistics AFOSR-88-0308 -444200-22523	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 90 1107
6a. NAME OF PERFORMING ORGANIZATION University of California, Davis / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NM
6c. ADDRESS (City, State, and ZIP Code) Divisicn of Statistics Univeristy of California Davis, CA 95616	7b. ADDRESS (City, State, and ZIP Code) Bolling Air Force Base Washington, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR / 8b. OFFICE SYMBOL (If applicable) NM	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR 88-0308
8c. ADDRESS (City, State, and ZIP Code) Bolling Air Force Base Washington, D.C. 20332-6448	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO
61102F	2304	A5	

11. TITLE (Include Security Classification)
Grant AFOSR 84-0159 Final Report - "Reliability Modeling and Inference for Coherent Systems Subject to Aging, Shock or Repair"

12 PERSONAL AUTHOR(S)
Francisco J. Samaniego

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 880801 TO 900731	900831	13

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This final report on grant AFOSR 88-0308 describes research accomplished during the period August 1, 1988-July 31, 1990 by Principal Investigator F.J. Samaniego and his collaborators under grant support. The results obtained in 8 completed manuscripts and two manuscripts in preparation are summarized. This research includes contributions to (i) product moment computation for multivariate survival functions (ii) parametric and nonparametric estimation in reliability (iii) the foundations of statistical estimation theory and (iv) nonstandard sampling techniques in life testing experiments.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21 ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	UNCLASSFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Jon A. Sjogren	22b TELEPHONE (Include Area Code) (202) 767-4940 / 22c. OFFICE SYMBOL AFOSR/NM

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

Final Report

AFOSR Grant 88-0308

Francisco J. Samaniego, Principal Investigator

August 31, 1990

During the period August 1, 1988 - July 31, 1990, the principal investigator and his co-investigators published, or submitted for publication seven research papers on the general topic "Reliability Modeling and Inference for Coherent Systems Subject to Aging, Shock on Repair" with the support of grant AFOSR 88-0308. This work is briefly described below.

1. "Estimating a Survival Curve when New is Better Than Used in Expectation," Journal of Naval Research Logistics, 36(1989), 693-708, (with L.R. Whitaker).

Summary: Let X be a positive random variable. The distribution F of X is said to be "new better than used in expectation," i.e., NBUE, if $E(X) \geq E(X-t|X>t)$ for all $t \geq 0$. Suppose $X_1,\ldots,X_n$, is a random sample from an NBUE distribution F. The problem of estimating F by a distribution which is itself NBUE is considered. The estimator G_n, defined as the NBUE distribution supported on the sample which minimizes the (sup norm) distance between the NBUE class and the empirical distribution function, is studied. The strong uniform consistency of G_n is proven, and a numerical algorithm for obtaining G_n is given.

2. "New Moment Identities Based on the Integrated Survival Function," IEEE Transactions on Reliability, Vol. R-38 (1989), 358-61, (with B. Arnold and D. Reneau).

Summary: The identity $E\{X\} = \int_0^\infty \bar{F}(x)dx$, where X is a nonnegative random variable and $\bar{F}(x) = Pr\{X>x\}$, is well known in reliability theory and survival analysis. In a 1982 paper, Samaniego extended this identity to produce an expression for the kth moment of a nonnegative random variable in terms of a multiple integral involving the survival function:

$$E\{X^k\} = k! \int_0^\infty \cdots \int_0^\infty \bar{F}(x_1 + \cdots + x_k)dx_1 \cdots dx_k$$

when the expectation exists. This identity can be used in establishing moment inequalities for the New Better Than Used (NBU) class of distributions. In this paper, we generalize the identity above to arbitrary multivariate random vectors. This result is then used to derive inequalities for multivariate extensions of the NBU family of distributions.

3. "On Estimating the Reliability of Systems Subject to Imperfect Repair," Journal of the American Statistical Association, 84(1989), 301-309, (with L.R. Whitaker).

Summary: This study of statistical inference for repairable systems focuses on the development of estimation procedures for the life distribution F of a new system based on data on system lifetimes between consecutive repairs. The Brown-Proschan 'imperfect repair' model postulates that, at failure, the system is repaired to a condition as good as new with probability p and is otherwise repaired to the condition just prior to failure. In treating issues of statistical inference for this model, we begin by pointing out the lack of identifiability of the pair (p,F) as an index of the distribution of interfailure times $T_1, T_2, \ldots$ We show that data pairs (T_i, Z_i), i=1,2... render the parameter pair (p,F) identifiable, where Z_i is a Bernoulli variable which records the mode of repair (perfect or imperfect) following the ith failure. Under the assumption that data of the form $\{(T_i, Z_i)\}$ are drawn via inverse sampling until the occurence of the m^{th} perfect repair, we study the problem of estimating the parameter pair (p,F) of the Brown-Proschan model. We demonstrate that the nonparametric maximum likelihood estimator of F exists only in special cases, but that a neighborhood maximum likelihood estimator $\hat{F}$ (using the language of Kiefer and Wolfowitz (1956)) always exists and may be derived in closed form. Under mild assumptions, we demonstrate the strong uniform consistency of $\hat{F}$ and the weak convergence of an appropriately scaled version of $\hat{F}$ to a Gaussian process. It is noted that these results apply to other experimental designs, such as renewal testing, and that they can be extended to the age-dependent imperfect repair model of Block, Borges and Savits (1985).

4. "Estimating a Survival Curve When New is Better Than Used of a Specified Age," Journal of the American Statistical Association. 85 (1990), 123-31, (with D. Reneau).

Summary: Suppose the survival function S of a random variable X is a member of the NBU-$\{t_o\}$ class. As defined by Hollander, Park, and

Proschan (1986), S must satisfy the condition $S(x+t_o) \leq S(x)S(t_o)$ for all $x \geq 0$. Given a random sample from S, an estimator is sought that itself belongs to the NBU-$\{t_o\}$ class and has the same asymptotic optimality properties as the empirical survival curve S_n. We study the estimator $\hat{S}$ of S given implicitly by $\hat{S}(x) = \min [S_n(x), \hat{S}(t_o)\, \hat{S}(x-t_o)]$ when $x > t_o$, with $\hat{S}(x) = S_n(x)$ for $x \leq t_o$. We derive a closed-form representaton for $\hat{S}$ and show that it is a strong uniformly consistent NBU-$\{t_o\}$ estimator of S which converges to S at the best possible (pointwise and mean square) rates. The asymptotic distribution theory of $\hat{S}$ is discussed, and is derived explicitly for functions S for which $S(x+t_o) < S(x)S(t)$ for all $x \geq 0$. Confidence procedures based on $\hat{S}$ are also discussed. In the final section, we describe the results of a simulation study which lends strong support to the claim that $\hat{S}$ estimates an NBU-$\{t_o\}$ survival curve more precisely than does the empirical survival function S_n.

5. "Estimating a Survival Curve When New is Better Than Used with Respect to a Distinguised Set," Journal of Naval Research Logistics, to appear (with D. Reneau and R. Boyles).

Summary: Hollander, Park, and Proschan (1985, 1986) define a survival function S of a positive random variable X to be New Better Than Used at Age t_o (NBU-$\{t_o\}$) if S satifies $S(x+t_o) \leq S(x)\, S(t_o)$ for all $x \geq 0$, where $S(x) = P(X > x)$. The NBU-$\{t_o\}$ class is a special case of the NBU-A family of survival distributions, where A is a subset of $[0,\infty)$. These families introduce a variety of modeling posibilities for use in reliability studies. We treat problems of nonparametric estimation of survival functions from these classes by estimators which are themselves

members of the classes of interest. For a number of such classes, a recursive estimation technique is shown to produce closed-form estimators which are strongly consistent and converge to the true survival distribution at optimal rates. For other classes, additional assumptions are required to guarantee the consistency of recursive estimators. As an example of the latter case, we demonstrate the consistency of a recursive estimator for $S \in \text{NBU-}[t_0,\infty)$ based on lifetime data from items surviving a preliminary "burn-in" test. The relative precision of the empirical survival curve and several recursive estimators of S are investigated via simulation; the results provide support for the claim that recursive estimators are superior to the empirical survival curve in restricted nonparametric estimation problems of the type studied here.

6. "On Nonparametric Maximum Likelihood Estimation of a Distribution Uniformly Stochastically Smaller than a Standard," Statistics and Probability Letters, to appear (with J. Rojo).

Summary: The class of distributions which are uniformly stochastically smaller than a standard arises naturally in life testing experiments in which a system is tested in an environment known to be more severe or stressful than the environment in which the system has been pretested. The problem of estimating a distribution in this class via nonparametric maximum likelihood is considered. The NPMLE is derived in closed form, and its inconsistency is demonstrated. The form of the NPMLE suggests a possible avenue for developing consistent nonparametric estimators for this class.

7. "On Estimating a Survival Curve Subject to a Uniform Stochastic Ordering Constraint", submitted for publication (with J. Rojo).

Summary: If F, G are cummulative distribution functions on $[0,\infty)$ governing the lifetimes of specific systems under study, and $\bar{F}$, $\bar{G}$ are their corresponding survival functions, then F is said to be uniformly stochastically smaller than G, denoted by $F <_{(+)} G$, if and only if the ratio $\ell(x) = \bar{G}(x)/\bar{F}(x)$ is nondecreasing for $x \in [0, \sup\{t:\bar{F}(x) > 0\})$. When F and G are absolutely continuous, $F <_{(+)} G$ is equivalent to the assumption that the corresponding failure rates are ordered. The applicability of the notion of uniform stochastic ordering in reliability and life-testing is discussed. Given that a random sample $X_1,\ldots X_n$ of lifetimes has been obtained from F, where F is assumed to satisfy the uniform stochastic ordering constraint $F <_{(+)} G$ (or alternatively, $F >_{(+)} G$), where G is fixed and known, the problem of estimating F is addressed. While the method of nonparametric maximum likelihood estimation fails, in general, to provide consistent estimators in this type of problem, a recursive approach is shown to provide estimators which converge uniformly to F with probability one and are as close or closer to F, in the sup norm, as the empirical distribution function. This leads to a proof of the inadmissibility of the empirical distribution function, relative to the sup norm loss criterion, when estimating $F <_{(+)} G$ (or $F >_{(+)} G$) with G continuous. The two-sample estimation problem is also discussed.

In addition to the work in the manuscripts described above, work on a problem in the foundations of statistical estimation theory has been completed, and appears as Technical Report No.202, Division of Statistics, University of California, Davis. This work is summarized below:

8. "Towards a Reconciliation of the Bayesian and Frequentist Approaches to Estimation", submitted for publication (with D.M. Reneau).

Summary: The Bayesian and frequentist approaches to estimation are reviewed. The status of the debate regarding the use of one approach over the over is discussed, and its inconclusive character is noted. A criterion for comparing Bayesian and frequentist estimators within a given experimental framework is proposed. The competition between a Bayesian and a frequentist is viewed as a game with the following components: a random observable, a true prior distribution unknown to both statisticians, an operational prior used by the Bayesian, a fixed frequentist rule (say, the MVUE) used by the frequentist and a fixed loss criterion. This competition is studied in the context of exponential families, conjugate priors and squared error loss. The class of operational priors which yield Bayes estimators superior to the best frequentist estimator is characterized. The implications of the existence of a threshold separating the space of prior distributions into good and bad priors are explored, and its relevance in areas such as Bayesian robustness and the elicitation of prior distributions is discussed. Both the theoretical and empirical results of this article point to the following general conclusion: The Bayesians have been right all along! And so have the frequentists! Both are right (and better than the other) under specific (and complementary) circumstances.

Work has been completed on two additional problems, and technical reports on these problems are in preparation. The results to be presented in these upcoming reports are described below:

9. "Life testing under Ranked Set Sampling ", in preparation, (with P.H. Kvam).

A typical observation in a ranked set sample is an observation of the ith smallest item in a random sample of size n (with the other n-1 items in that random sample being unobserved). Let $X_{i:n:k}$ represent the outcome of the kth trial in which the ith smallest item from a random sample of size n is recorded. We will say that the ranked set sample is balanced if, for every i, the ith order statistic is independently observed the same number of times (say m times). A balanced ranked set sample may be displayed in a rectangular array such as the following:

$$\begin{array}{cccc}
X_{1:n:1} & X_{1:n:2} & \cdots & X_{1:n:m} \\
X_{2:n:1} & X_{2:n:2} & \cdots & X_{2:n:m} \\
\cdot & & & \\
\cdot & & & \\
\cdot & & & \\
\cdot & & & \\
X_{n:n:1} & X_{n:n:2} & \cdots & X_{n:n:m}
\end{array}$$

A ranked set sample arises in reliability experiments when observations consist of the lifetimes of k- out-of-n systems in iid components, where k may vary from system to system. In the balanced ranked set sampling scenario, the average $\bar{X}_{RSS}$ of the m.n observation has been studied as an estimator of the mean of the population (assuming all observations are obtained from the same underlying distribution). The mean of the ranked set sample has a number of good properties, among which is the fact that it is an unbiased estimator of the population mean and has variance no larger (and possibly much smaller) than that of the mean of a random sample of the same size. Stokes and Sager (JASA, 1988) recently proposed the empirical distribution of the values obtained in a ranked set sample as an estimator of the population distribution of function F. Moreover, in analogy to results concerning $\bar{X}$, they have shown that this estimator is unbiased and is more precise than the empirical distribution of a random sample of the same size.

The goal of our research in this area has been to construct improved estimators for the mean μ and for the underlying distribution function F under ranked set sampling. There is a need to develop estimators of μ and F for unbalanced ranked set samples, and to consider questions of optimality in both balanced and unbalanced data sets. In the present paper, we derive, in a variety of settings, the best linear unbiased estimators of μ, and show, as a byproduct, that $\bar{X}_{RSS}$ is generally inadmissible! We are able to produce estimators of μ in several parametric problems (including ranked set sampling from the exponential and the normal models) which are unbiased and are uniformly more precise than the standard estimator $\bar{X}_{RSS}$. Our examination of alternative nonparametric estimators of the underlying distribution F based on a ranked set sample has also led to estimators that are more precise in the balanced case and that extend naturally to the unbalanced case, retaining desirable fixed sample-size and asymptotic properties.

10. Life Testing in Variably Scaled Environments, in preparation (with P.H. Kvam).

Summary: Chapter 9 of the book Methods for Statistical Analysis of Reliability and Lifetime Data by Mann, Schafer and Singpurwalla present the classical theory of accelerated life testing. The motivation for these methods comes from the fact that tests of systems or components under normal operating conditions often yield very sparse data since only a few items will fail during the (relatively short) test period that is feasible in most experiments. One is thus led to consider testing components under higher stress than normal, and extrapolating from such experiments to obtain estimated performance under normal conditions. Extrapolation is always controversial, of course, and the best applications of existing methods of accelerated life testing tend to be those which make use of as much data as possible from low stress environments in conjuction with failure data obtained under high stress.

Classical methods of a accelerated life testing have a number of limitations. First, these methods are only applicable to a single component rather than several components (e.g., all components in a system) simultaneously. Secondly, these methods tend to require the specification of exact relationships between the parameters of the lifetime distribution at one stress level and the parameters at another stress level. An example of a popular model of this type is the "power rule", which postulates the relationship $\mu_i = c\ S^{-m}$, where μ_i is the mean

time to failure under stress S, and c and m are unknown positive parameters to be fit to data. The aforementioned text discusses several such functional relationships, including the well-known Arrhenius and Eyring models. The difficulty in using the classical ALT methods in real applications (and the primary source of their misuse) is the dubious validity of specific parameter/stress level relationships in the application at hand. There have been a handful of papers which study nonparametric models without specific stress-lifetime functional relationships assumed, but their attendant statistical analyses are, to date, mostly ad hoc.

In a 1986 paper in the Journal of Applied Probability, Lindley and Singpurwalla proposed a model for life testing of several components in multiple environments. The model is seemingly quite simple: the lifetime of component i in environment j is assumed to be independently exponentially distributed with failure rate $\lambda_i \cdot \theta_j$. One of the main points of their paper was to show how dependence of component lifetimes might arise. They derive the joint distribution of component lifetimes under specific models for random selection of the environment. While Lindley and Singpurwalla mention accelerated life testing in motivating their model, they do not empasize this application and, instead, concentrate on the impact of averaging over environmental affects which are assumed to be random.

The Lindley-Singpurwalla model is interesting and useful in that it permits the simultaneous treatment of several different components, and postulates only that the several environments under which these components are tested are different, via a scale change. The model does not require the specification of a functional relationship between stress and expected failure time; the absense of such restrictions allows one to consider applications such as life testing for aircraft operating in weather conditions which vary with the region in which the plane is based (an application where the exact "stress" attributable to particular weather conditions is difficult to quantify).

The L-S model is appropriate in multi-environment life testing when a change in environment has a similar affect on all components (e.g., halves the expected life). It's important, of course, to be able to test that aspect of the model when applying it to data. The model is flexible enough to accomodate reliability growth as well as degradation, and can be applied to situations where no prior assumptions are made regarding the relative harshness of the several environments involved.

While Lindley-Singpurwalla model is quite appealing, there's a price to pay for the positives above. It turns out that statistical inference for the L-S model poses some rather difficult technical problems. Until very recenlty, for example, it looked like the problem of deriving maximum likelihood estimates of model parameters was completely intractable. We are now able to provide a relatively complete solution of that problem. Although the problem of deriving the MLE from data on k components in r environments (i.e., the k + r - 1 parameter problem) does not lend itself to closed-form solutions, we have been able to establish the existence and uniqueness of a solution, and have proven the local concavity of the likelihood surface is a neighborhood of the MLE. This latter fact will guarantee the convergence of the numerical method we use (the Gauss-Sidel method) given an appropriate set of initial values. The present paper establishes the statistical feasibility of the use of Lindley-Singpurwalla model in accelerated life testing. It also provides the necessary tool for testing the validity of model. Our results make it possible to carry out a likelihood ratio test of the hypothesis that environments affect components uniformly and are related to one another via scale change.

Finally, under the assumption that the Lindley-Singpurwalla model holds, we have studied the relative efficiency of estimates of the baseline (normal use) failure rates $\{\lambda_i\}$ with and without the data obtained from tests in other environments. When experimentation is done in r environments, and the ratio of a sample size in the normal-use environment to the sample size in other environments is $p \in (0,1)$, the relative efficiency of our failure rate estimates is approximately $(t-1+p)/p$. For $t = 5$ and $p = .2$, for example, our estimators are more than 20 times more efficient than estimators based on normal use data alone.